

On Average Distortion of Embedding Metrics into the Line

Yuri Rabinovich

Received: 20 October 2003 / Published online: 5 February 2008
© Springer Science+Business Media, LLC 2008

Abstract We introduce and study the notion of the *average distortion* of a nonexpanding embedding of one metric space into another. Less sensitive than the multiplicative metric distortion, the average distortion captures well the global picture and, overall, is a quite interesting new measure of metric proximity, related to the concentration of measure phenomenon. The paper mostly deals with embeddings into the real line with a low (as much as it is possible) average distortion. Our main technical contribution is that the shortest-path metrics of special (e.g., planar, bounded treewidth, etc.) undirected weighted graphs can be embedded into the line with *constant* average distortion. This has implications, e.g., on the value of the MinCut–MaxFlow gap in uniform-demand multicommodity flows on such graphs.

Keywords Metric embeddings · Average distortion

1 Introduction

The theory of finite metric spaces is an attractive young branch of discrete mathematics and theoretical CS. An overview of some of its methods, results and striking applications appears, e.g., in recent but already somewhat dated literature [9, 14]. Numerous more recent results appear in the Proceedings of *STOC'03* and *FOCS'03*. (After the appearance of this paper further spectacular results were obtained; see, e.g., [1, 5, 10, 12].)

The key role in these developments is played by the notion of *proximity* between metric spaces or, using other words, the *faithfulness of representation* of one metric

A preliminary version of this paper has appeared at *STOC'03*, pp. 156–162.

Supported in part by a grant ISF-247-02-10.5.

Y. Rabinovich (✉)

Department of Computer Science, University of Haifa, Haifa 31905, Israel
e-mail: yuri@cs.haifa.ac.il

space by another. One particularly successful (but by no means unique) measure of proximity is the *multiplicative distortion*. The multiplicative distortion of embedding a given finite metric space (X, d) into the host space (H, δ) can be defined as the minimum possible value of *contraction* of a nonexpanding embedding $f : X \rightarrow H$. The contraction of f is $\max_{x,y \in X} d(x, y)/\delta(f(x), f(y))$. The smaller the distortion, the more faithful the embedding.

In this paper we introduce and study a different (yet related) measure of proximity, which we call the *average distortion*. It is defined as the minimum possible value of $\sum_{x,y} d(x, y)/\sum_{x,y} \delta(f(x), f(y))$ over all nonexpanding $f : X \rightarrow H$, i.e., the ratio of the average distances. We are particularly interested in the case when the host space is \mathbb{R} , i.e., the line.

The average distortion of embedding a metric space into the line is a new metric-theoretic parameter. We argue that there are good reasons for studying it. Here are some of them. First, it is a very natural parameter and, without being explicitly named, it has already been seriously studied in modern convex geometry in the context of the concentration of measure phenomenon (see e.g., [6, 14, 16]). These studies have led to deep and surprising findings, which can safely be called fundamental. Second, this parameter turns out to be related to the quality of performance of the best currently known approximation algorithm for the sparsest cut problem. The (approximately) sparsest cut is used in countless algorithms involving graphs and networks, and is of prime importance in the area of algorithmic design. Third, we show that this parameter is capable of distinguishing special metric spaces from the general ones. One of the more interesting directions of study in the theory of finite metric spaces are the relations between the topology of the underlying graph, and the structure of its shortest-path metrics (under all nonnegative edge-weightings). For example, one of the better known open questions about the shortest-path metrics of special (planar, bounded treewidth, etc.) graphs is whether they can be embedded into ℓ_1 with a *constant* multiplicative distortion. Despite some “supporting evidence”, and a positive answer for all series-parallel and k -outerplanar graphs (see [3, 7]), the general question remains widely open. The main technical contribution of this paper can be viewed as a “cousin” of this conjecture. We show that the shortest-path metrics of such graphs can be embedded into the *line* with a *constant* average distortion. This surprising result does not hold for general metric spaces, where the average distortion can be as large as $\Omega(\log n)$, where $n = |X|$. When (X, d) is a submetric of a Euclidean space, which is a rather severe restriction, the average distortion can still be as large as $\Omega(\sqrt{\log n / \log \log n})$ (see [15]).

In general, the average distortion of embedding (X, d) into the line is bounded from above by $c_1^{\text{dom}}(d)$, the multiplicative distortion of approximating d by a convex combination of d -dominated line metrics. The latter parameter was studied in [2, 15]. It is at most $O(\log n)$ for general metrics, and at most $O(\sqrt{\log n})$ for the Euclidean ones. These bounds are tight, and coincide with those for average distortion. For planar metrics, however, $c_1^{\text{dom}}(d)$ can be as large as $\Omega(\sqrt{\log n})$ (as implied by [17]), whereas the average distortion is always constant.

The paper is organized as follows. We start with a short discussion of the metric concentration of measure phenomenon, and its relations to the average distortion. Next, we shortly discuss the main result of [11], a backbone of numerous algorithms for planar and other special graphs, and explain that in fact it is equivalent to the claim

that the average distortion of embedding shortest-path metrics of special graphs into ℓ_1^∞ is constant. In the second (the main) part of the paper we focus on embeddings into the line, establish basic general inequalities and bring some examples, and then prove the main result of this paper, i.e., that the average distortion of embedding the shortest-path metrics of special graphs into the line is constant. While, in view of the above, this is a considerable strengthening of the main result of [11], our proof is to a large extent based on the same KPR lemma, and its additional properties as established in [18]. We conclude the paper with further sharpening the main result for graphs of a constant treewidth, using a different proof technique.

A remark on our definition of the average distortion is due. In this paper, the average distortion is the minimum of $\sum_{x,y} d(x,y)/\sum_{x,y} \delta(f(x), f(y))$ over all nonexpanding f . This definition is by no means the only possible and, for instance, the minimum of $\sum_{x,y} \delta(f(x), f(y))/\sum_{x,y} d(x,y)$ over all noncontracting f is as natural, and at the first glance as interesting, as the one we use. But only at the first glance. For embeddings into the line, the first definition leads to average distortion $O(\log n)$ in the worst case, with fine relations between the metric structure and the corresponding value of distortion. The second definition leads to $O(n)$, which is tight already for the uniform metric, the simplest metric of all. The contrast is even more striking if we consider embeddings into trees. Corollary 2.9 of the current paper implies that, under the first definition, the average distortion of embedding d into a tree (and, more generally, into any special graph) is equal up to a constant multiplicative factor to the average distortion of embedding it into the line. Under the second definition, however, any metric can be embedded into a *star* so that the average distortion is less than 2 (see, e.g. [8], Lemma 14.7.1).

1.1 Preliminaries

Metrics Throughout the paper, the metric spaces are assumed to have the same universal underlying set $[n]$ (unless indicated otherwise). Thus, speaking of a metric d , we usually have in mind the metric space $([n], d)$. No distinction is made between metrics and semimetrics, i.e., two different points may coincide.

The *average distance* of d is $\text{av}(d) = 1/n^2 \cdot \sum_{x,y \in [n]} d(x,y)$. Notice that the sum is taken over the *ordered* pairs of points.

A *nonexpanding embedding* of a metric d into a host space (Y, δ) is a map $f : [n] \rightarrow Y$ such that the induced submetric d' of Y is dominated by d , i.e., no distance increases. Formally, for all $i, j \in [n]$, it holds $d'(i, j) = \delta(f(i), f(j)) \leq d(i, j)$. The key term of this paper is $\text{av}_Y(d)$, the maximum of $\text{av}(d')$ over all such d' . In particular, we shall be interested in $\text{av}_{\text{line}}(d)$ and $\text{av}_1(d)$, where the host spaces are \mathbb{R} and ℓ_1^∞ , respectively. The *average distortion* of embedding d into Y is defined as the ratio $\text{av}(d)/\text{av}_Y(d)$. It is always ≥ 1 , with equality iff f is an isometry.

The *median* of the metric d on $[n]$ is the point $p \in [n]$ minimizing the expression $1/n \sum_{x \in [n]} d(p, x)$; it will be denoted by med . Observe that for a set of points on a line, the order median coincides with the metric median. The following basic fact will be frequently used:

$$\frac{1}{n} \sum_{x \in [n]} d(\text{med}, x) \leq \text{av}(d) \leq \frac{2}{n} \sum_{x \in [n]} d(\text{med}, x). \quad (1)$$

The first inequality simply says that the average (with respect to p) value of $1/n \cdot \sum_{x \in [n]} d(p, x)$ is no less than its minimum value. The second inequality holds for any $p \in [n]$, since

$$\sum_{x,y \in [n]} d(x, y) \leq \sum_{x,y \in [n]} d(x, p) + d(p, y) = 2n \sum_{x \in [n]} d(p, x). \tag{2}$$

For a set $S \subseteq [n]$, let $\text{diam}(S) = \max_{x,y \in S} d(x, y)$ be the d -diameter of S . Let also $\xi(S)$ be the average d -distance to S , $\xi(S) = 1/n \cdot \sum_{x \in [n]} d(x, S)$, where $d(x, S) = \min_{y \in S} d(x, y)$. Moreover, define

$$\text{diam}_c(d) = \min_{S \subseteq [n], |S| = \lceil cn \rceil} \text{diam}(S); \quad \xi_c(d) = \max_{S \subseteq [n], |S| = \lceil cn \rceil} \xi(S).$$

The shortest-path metric (and its submetrics) of a weighted planar, bounded treewidth, etc., graph will be called *planar, bounded treewidth, etc.*, respectively. The weights, of course, must be nonnegative. A *cut metric* δ_S on $[n]$ for $S \subseteq [n]$ has value 1 if x, y are separated by S , and 0 otherwise. An important fact about the ℓ_1 -embeddable metrics on $[n]$ is that they are precisely the metrics which can be represented as a nonnegative combination of cut metrics on $[n]$.

Graphs The graphs are assumed to be connected and undirected, with nonnegative weights on the edges. They will often be regarded as solid one-dimensional geometric objects, the edges being intervals of length $d(e)$ glued at the vertices. The shortest-path metric d of G naturally extends (in a linear manner) to the entire solid G . The vertex set of G will usually be $[n]$. We shall call a graph *special* if it excludes a specified fixed minor M . In particular, such are the planar graphs and the graphs of bounded treewidth.

The *vertex expansion* Φ^v of G is defined as

$$\Phi^v = \min_{S \subseteq V, |S| \leq n/2} |\partial^v S| / |S|,$$

where $\partial^v S$ is the vertex boundary of S , i.e., the set of all vertices in \bar{S} which are adjacent to S .

In addition to the weights, edges may have nonnegative *capacities* $C = \{c_e\}_{e \in E[G]}$. For $S \subseteq [n]$, denote by $|\partial_C S|$ the sum of the C -values of all the edges crossing from S to \bar{S} . The *sparsest cut* is defined as

$$T_C = \min_{S \subseteq [n]} \frac{|\partial_C S|}{|S| \cdot |\bar{S}|}.$$

1.2 Average Distortion and the Metric Concentration of Measure

When a metric space (X, d) is embedded into a poorer metric space (H, δ) in a non-expanding manner, it often happens that X collapses, i.e., that all but an insignificant portion of its image (necessarily) lies in a ball of a (relatively) tiny radius. Moreover, the density of the image is rapidly declining as one gets away from this ball. This phenomenon is called the concentration of measure; it is of fundamental importance

in modern convex geometry and probability theory (see, e.g., [6, 14, 16]). As an illustration, consider a nonexpanding embedding f of C_n , the n -cube $\{0, 1\}^n$ equipped with the Hamming metric, into \mathbb{R} . Let $\text{med}(f)$ be the median of the multiset $f(C^n)$. Paul Levy’s inequality for C_n (see, e.g., [14], Theorems 14.2.3, 14.3.2) claims that for any $c > 0$ it holds (with respect to the uniform probability measure on the vertices):

$$\Pr[f(x) - \text{med}(f) \geq cn^{0.5}] \leq e^{-c^2/2}. \tag{3}$$

The same, of course, applies to the distribution of $\text{med}(f) - f(x)$. This means that almost the entire image of C_n must lie in an interval of length $O(n^{0.5})$, and that the rate of density drop is $e^{-c^2/2}$, where the distance from the median is measured in units of $n^{0.5}$. Thus, the n -cube collapses, and the “typical” distance drops from $n/2$ to $O(n^{0.5})$. Notice that for a particular nonexpanding function $f(x) = \sum_{i=1}^n x_i$, (3) becomes the familiar Chernoff bound.

What is the “typical” distance? In this paper, rather than looking at the radius of the minimal ball containing a significant part of the image of (X, d) , we look at the average distance in the image, and compare it to the original average distance. While gaining in simplicity and in applicability, we lose little in the original flavour. The average distortion can thus be viewed as an attempt to quantify the measure concentration.

The drop rate, the other important parameter associated with concentration of measure, is closely related to the isoperimetrical properties of the metric. Its definition is meaningful for homogeneous (in particular, vertex-transitive) metric spaces; it is not entirely clear how to extend it to nonhomogeneous ones. Perhaps the drop rate should be considered locally. In the present paper we shall not attempt to study it. Let us just remark that when it is well defined, it is extremely useful, and demonstrate this by an example.

Claim 1.1 *A lower bound on the dimension $r = r(\gamma)$ so that the Hamming metric on the n -cube can be embedded into ℓ_∞^r with average distortion γ , is $r = (2^n)^{\Omega(1/\gamma^2)}$.*

Proof Let d be the Hamming metric, let f be a nonexpanding embedding of d into ℓ_∞^r , and let d' be the resulting d -dominated induced submetric of f . Finally, let $\{d'_i\}_{i=1}^r$ be the d -dominated line metrics induced by the corresponding coordinate functions $\{f_i\}_{i=1}^r$.

Let $z > 0$ be a value to be specified later. Interpreting $\text{av}(d')$ as the expectation with respect to the uniform distribution on the unordered pairs of vertices, we get

$$\begin{aligned} \text{av}(d') &= \mathbb{E}[d'(x, y)] = \int_0^\infty \Pr[d'(x, y) \geq u] \, du \\ &\leq z \cdot m^{0.5} + \int_{z \cdot m^{0.5}}^\infty \Pr[d'(x, y) \geq u] \, du \\ &\leq z \cdot m^{0.5} + \sum_{i=0}^r \int_{z \cdot m^{0.5}}^\infty \Pr[d'_i(x, y) \geq u] \, du \\ &= z \cdot m^{0.5} + m^{0.5} \cdot \sum_{i=0}^r \int_z^\infty \Pr[d'_i(x, y) \geq tm^{0.5}] \, dt. \end{aligned}$$

Now, if $d'_i(x, y) \geq tm^{0.5}$, then at least one of x, y is at least $0.5 \cdot tm^{0.5}$ away from the corresponding $\text{med}(f_i)$. Plugging in the bound from (3), and setting $z = \sqrt{8 \ln r}$, we conclude that

$$\text{av}(d') \leq m^{0.5} \cdot \left(z + r \int_z^\infty 4e^{-t^2/8} dt \right) \leq (\sqrt{8 \ln r} + 4) \cdot m^{0.5}.$$

Thus, in order for $\gamma \cdot \text{av}(d') \geq \text{av}(d)$ to hold, it must hold $\gamma \cdot (8 \ln r^{0.5} + 4) \cdot m^{0.5} \geq m/2$, and the lower bound on r follows. \square

A similar analysis applied to the shortest-path metric of a constant-degree expander of size n yields an even stronger bound: $r(\gamma) = n^{\Omega(1/\gamma)}$. Instead of (3), one should use

$$\Pr[f(x) - \text{med}(f) > c] \leq 0.5 \cdot (1 + \Phi^v)^{-\lfloor c \rfloor}, \tag{4}$$

where Φ^v is the vertex expansion of the graph. This bound follows from the following simple argument. Observe that for any $c \in \mathbb{N}$, and any subset R of the vertices of size $\geq n/2$, it holds, by definition of Φ^v ,

$$|\{x, d(x, R) > c\}| \cdot (1 + \Phi^v) \leq |\{x, d(x, R) > c - 1\}|,$$

implying that for any $c \geq 0$,

$$|\{x, d(x, R) > c\}| \cdot (1 + \Phi^v)^{\lfloor c \rfloor} \leq |\bar{R}| \leq 0.5n.$$

Defining $R = \{x | f(x) \leq \text{med}(f)\}$, and using the fact that f is nonexpansive, we arrive at (4).

1.3 Average Distortion and the Main Result of [11]

Let G be a graph with nonnegative capacities $C = \{c_e\}$ on the edges. Consider the form

$$F(D) = \frac{\sum_E c_e D(e)}{\sum_{x,y} D(x, y)}.$$

As observed in [13], $\min_{\delta \in \ell_1} F(\delta)$ is attained at some cut metric δ_S , and therefore is equal to a half of T_C , the sparsest cut of G . On the other hand, $\min_d F(d)$, where the minimum ranges over all metrics on $[n]$, is attained at a shortest-path metric of G with some edge weights, which can be efficiently computed. The main result of [11] is a polynomial-time constant-factor approximation algorithm for the sparsest cut of special graphs, based on the following statement:

Theorem 1.2 *For any shortest-path metric d of G with a forbidden minor of size r there exists, and can be found in polynomial time, a subset of vertices $S \subset [n]$, for which $F(\delta_S)/F(d)$ is at most $O(r^3)$.*

This theorem turns out to be completely equivalent to the following statement about the average distortion of embedding such d into ℓ_1 :

Theorem 1.3 *There exists, and can be found in polynomial time, a d -dominated nonnegative combination d' of cut-metrics $\{\delta_S\}$, such that $\text{av}(d') \geq \Omega(r^{-3}) \cdot \text{av}(d)$. In particular, the average distortion of embedding d into ℓ_1 is $O(r^3)$.*

The implication Theorem 1.3 \Rightarrow Theorem 1.2 is quite simple; it follows from a standard argument first employed in [13]. Representing d' as a nonnegative combination of polynomially many cut metrics $\{\delta_S\}$, and trying them all, one finds δ_S for which $F(\delta_S) \leq F(d')$. The corresponding S satisfies the desired requirements, since

$$\begin{aligned} F(\delta_S) \leq F(d') &= \frac{\sum_E c_e d'(e)}{\sum_{x,y} d'(x,y)} \leq \frac{\sum_E c_e d(e)}{\sum_{x,y} d(x,y)} \\ &\leq \frac{\text{av}(d)}{\text{av}(d')} \cdot \frac{\sum_E c_e d(e)}{\sum_{x,y} d(x,y)} = O(r^3) \cdot F(d). \end{aligned}$$

The other direction, Theorem 1.2 \Rightarrow Theorem 1.3, is based on the strong LP duality, and is somewhat technical. Since Theorem 1.3 is considerably weaker than the central result of this paper, Theorem 2.5 (as ℓ_1^∞ is considerably richer than the line), we prefer to omit here the proof of the former.¹

2 Embeddings into the Line

2.1 Basic Inequalities and Examples

In order to bring some interesting examples of the behaviour of $\text{av}_{\text{line}}(d)$, let us first show the approximate identity between $\xi_c(d)$, $c \in (0, 0.5]$, and $\text{av}_{\text{line}}(d)$:

Proposition 2.1 $4\xi_c(d) \geq \text{av}_{\text{line}}(d) \geq 2c\xi_c(d)$.

Proof We start with the first inequality. Observe that $\xi_c(d)$ is monotone nonincreasing in c , and thus it suffices to deal with $c = 0.5$. Consider the d -dominated line metric (X, d') for which $\text{av}(d') = \text{av}_{\text{line}}(d)$, and let med' be the median of X . Let L and R be the subsets of the underlying set of d' lying (strictly) to the left and to the right of med' , respectively, and let M denote the remaining points (i.e., those coinciding with med'). Assuming w.l.o.g., that $\sum_{x \in L} d'(\text{med}', x) \leq \sum_{x \in R} d'(\text{med}', x)$, we augment R with a subset of points from M to get a set S of size $\lceil n/2 \rceil$. In particular, $\text{med}' \in S$. By (2) from Sect. 1.1,

$$\begin{aligned} \text{av}_{\text{line}}(d) = \text{av}(d') &\leq \frac{2}{n} \sum d'(\text{med}', x) \leq 2 \cdot \frac{2}{n} \sum_{x \in S} d'(\text{med}', x) \\ &= \frac{4}{n} \sum_{x \in S} d(x, S) \leq 4\xi_{1/2}(d). \end{aligned}$$

¹The interested reader can find an outline of the equivalence proof and a discussion of the original proof of Theorem 1.2 in the preliminary version of this paper, in Proceedings of *STOC'03*.

For the second inequality, let $R \subset [n]$ be the set on which $\xi_c(d)$ is attained. Consider the following embedding f of d into the line: $f : x \rightarrow d(x, R)$, the induced line metric d' . This is clearly a nonexpanding embedding, and

$$\text{av}_{\text{line}}(d) \geq \text{av}(d') \geq \frac{1}{n^2} \cdot 2 \cdot \sum_{x \in R, y \notin R} d'(x, y) \geq \frac{2 \cdot |R| \cdot n}{n^2} \cdot \xi_c(d) = 2c\xi_c(d). \quad \square$$

Proposition 2.1 can be used for deriving upper bounds on $\text{av}_{\text{line}}(d)$ in the following manner. Observe that

$$\xi_{1/2}(d) = \frac{1}{n} \int_0^\infty |\{x, d(x, R) \geq t\}| dt,$$

where R is the set on which $\xi_{1/2}(d)$ is attained. Therefore, if d has nice isoperimetric properties, they can be used to obtain a nontrivial bound on $\xi_{1/2}(d)$, and hence on $\text{av}_{\text{line}}(d)$. In particular,

Claim 2.2 *Let G be a unit-weighted graph on $[n]$ with vertex expansion Φ^v , and let d be its shortest-path metric. Then, $\text{av}_{\text{line}}(d) = O(1/\Phi^v)$.*

Proof Let R be as before. By the discussion following (4), for any $t \geq 0$,

$$|\{x, d(x, R) \geq t\}| \cdot (1 + \Phi^v)^{\lfloor t \rfloor} \leq |\bar{R}|.$$

Hence,

$$\xi_{1/2}(d) \leq |\bar{R}|/n \int_0^\infty (1 + \Phi^v)^{-\lfloor t \rfloor} dt = O(1/\Phi^v). \quad \square$$

Claim 2.2 implies that for the shortest-path metric of a unit-weighted constant-degree expander, $\text{av}_{\text{line}}(d) = O(1)$, while $\text{av}(d) = \Theta(\log n)$. This well-known yet amazing fact remains true for $\text{av}_1(d)$ as well.

As another example, consider the Hamming Cube, C_n . By Harper’s isoperimetric inequality, $\Phi^v(C_n) = \Theta(1/\sqrt{n})$, implying that $\text{av}_{\text{line}}(C_n) = O(\sqrt{n})$, whereas $\text{av}_1(d) = \text{av}(d) = n/2$.

The following simple but important inequality allows us to establish the relationship between $\text{av}_{\text{line}}(d)$, $\text{diam}_c(d)$ and $\text{av}(d)$.

Proposition 2.3 *For any $R \subseteq [n]$,*

$$\text{diam}(R) + \frac{n}{|R|} \cdot \text{av}_{\text{line}}(d) \geq \text{av}(d).$$

Proof Observe first that

$$\sum_{x,y} d(x, y) \leq \sum_{x,y} [d(x, R) + d(y, R) + \text{diam}(R)] = 2n \sum_x d(x, R) + n^2 \text{diam}(R).$$

Consider now a nonexpanding embedding f of d into the line, where $f : x \rightarrow d(x, R)$, and let d' be the corresponding d -dominated line metric. Then,

$$\begin{aligned} \sum_{x,y} d'(x, y) &= \sum_{x,y} |d'(x, R) - d'(y, R)| \geq 2 \cdot \sum_{x \notin R, y \in R} |d'(x, R) - d'(y, R)| \\ &= 2|R| \cdot \sum_x d(x, R). \end{aligned}$$

This, together with the first inequality, implies the proposition. \square

Corollary 2.4 $\text{diam}_c(d) + 1/c \cdot \text{av}_{\text{line}}(d) \geq \text{av}(d)$.

2.2 Special Metrics: General Case

As mentioned in the Introduction, the construction of Bourgain implies that the gap between $\text{av}(d)$ and $\text{av}_{\text{line}}(d)$ never exceeds $O(\log n)$, and is attained, e.g., at the shortest-path metric of a constant degree expander. It is also known that for Euclidean metrics the gap is smaller: at most $O(\sqrt{\log n})$, with an almost matching lower bound of $\Omega(\sqrt{\log n / \log \log n})$ (see [15]). In this section we present the main technical result of this paper, and show that for a shortest-path metric of a special graph, the gap is constant!

Theorem 2.5 *Let G be a weighted graph lacking a minor M of size r , and let d be its shortest-path metric. Then,*

$$\text{av}_{\text{line}}(d) \geq \Omega(r^{-3}) \cdot \text{av}(d).$$

Proof The key ingredient of the proof is the following construction, known as the Klein–Plotkin–Rao Lemma [11]. Let G be a solid graph lacking the minor $K_{r,r}$, and let $\Delta > 0$ be a parameter.

In the first round, choose an arbitrary point p in the solid G , and choose a value t randomly and uniformly from $[0, \Delta)$. Remove from G all the points q for which $d(p, q) \equiv t \pmod{\Delta}$. As a result, the graph falls apart into connected components (regions). In the second round, do the same independently for each region obtained in the previous round. Continue in the same manner for r rounds. Output the obtained regions.

Lemma 2.6 *The construction has the following properties:*

1. *The d -diameter of each output region is $\leq 4r^2 \Delta$.*
2. *For each edge e , $\Pr[e \text{ is cut}] \leq r \cdot d(e) / \Delta$.*
3. *Let Q be set of all the removed points (equivalently, the boundary of the output regions). Call a point $x \in G$ good if $d(x, Q) \geq \frac{1}{2(r+1)} \Delta$. Then, for each $x \in G$, the probability that x is good is $\geq 1/e$.*

Properties 1 and 2 are from [11]; property 3 is from [18]. In this paper, we shall need only properties 1 and 3. Let A be the set of all good vertices. By property 3 of

construction, $E[|A|] \geq n/e$. Assume that A is indeed that large. The lemma is stated for graphs with forbidden minor $K_{r,r}$, however, since $K_{r,r}$ contains K_r as a minor, it applies to graphs lacking an arbitrary minor M of size r .

Apply the KPR construction with $\Delta = c_1 \cdot \text{av}(d)$, where $c_1 \in [0, 1]$ is a constant to be specified later. Let $c_2 \in [0, 1]$ be another constant, also to be specified later. There are two cases to consider.

Case 1. No Region S_i Contains More than c_2n Vertices For each S_i choose $\epsilon_i = \pm 1$ randomly and independently from the others. Define a mapping $f : V \rightarrow \mathbb{R}$ by $f(x) = \epsilon_i \cdot d(x, Q)$, where $x \in S_i$, and let d' be the induced line metric. Observe that d dominates d' , as follows from the triangle inequality for x, y lying in the same region, and from the fact that any path (and in particular the shortest one) between x and y lying in different regions must cross Q .

We claim that the expected value of $\text{av}(d')$ under the random choice of ϵ_i 's is sufficiently large. Indeed, if x and y are vertices lying in different regions, and at least one of them is good, then $E[d'(x, y)] \geq \frac{1}{2} \cdot \frac{1}{2(r+1)} \cdot \Delta$. Call such a pair of vertices *significant*.

$$E[\text{av}(d')] \geq \frac{\Delta}{4(r+1)} \cdot \frac{|\{(x, y) | (x, y) \text{ is a significant pair}\}|}{n^2}.$$

How many significant pairs are there?

$$\sum_i |A \cap S_i| \cdot (n - |S_i|) = |A| \cdot n - \sum_i |A \cap S_i| \cdot |S_i| \geq |A| \cdot n - |A| \cdot \max_i |S_i|.$$

By our assumptions there are at least n/e good vertices, and no S_i contains more than c_2n vertices. Thus, there are at least $e^{-1}(1 - c_2)n^2$ significant pairs. Consequently,

$$E[\text{av}(d')] \geq \frac{1 - c_2}{4e} \cdot \frac{1}{(r+1)} \cdot \Delta = \frac{c_1(1 - c_2)}{4e} \cdot \frac{1}{(r+1)} \cdot \text{av}(d).$$

Case 2. Some Region S_i Contains More than c_2n Vertices In this case, we use a mapping $f : V \rightarrow \mathbb{R}$ defined by $f(x) = d(x, S_i)$. By property 1 of the construction, $\text{diam}(S_i) \leq 4r^2\Delta$. Arguing as in the proof of Proposition 2.3, and keeping in mind the choice of Δ ,

$$\text{av}(d') \geq c_2 \cdot (\text{av}(d) - 4r^2\Delta) = c_2 \cdot (1 - 4r^2c_1) \cdot \text{av}(d).$$

It remains to specify the constants c_1 and c_2 . Setting $c_1 \approx \theta(r^{-2})$ and $c_2 \approx \theta(1)$, we arrive at the desired conclusion. □

The above proof can be easily turned into a deterministic polynomial-time algorithm. In order to ensure that the size of A is at least n/e , one can modify each step of [11] as follows. Instead of choosing the value t randomly, map all the vertices to the interval $[0, \Delta)$ by $v \rightarrow d(v, p) \bmod \Delta$, and find t for which the interval $[t - \frac{\Delta}{2(r+1)}, t + \frac{\Delta}{2(r+1)}) \bmod \Delta$ contains no more than $1/(r+1)$ fraction of the images. This will ensure that after r rounds there are at least $n \cdot (1 - \frac{1}{r+1})^r \geq n/e$ good

vertices. To eliminate the randomness altogether, observe that in the choice of ϵ_i 's it is sufficient to require only *pairwise* independence; therefore, applying a standard derandomization technique, a good choice can be made deterministically.

Remark 2.7 Often the term (e.g.) *planar metric* is used for an induced *submetric* on a subset $U \subseteq V[G]$, rather than on the whole $V[G]$. We would like to point out that Theorem 2.5 applies in this case as well; the only modification to the proof is that one should consider only the vertices in U and, respectively, n should denote the size of U instead that of V .

Remark 2.8 In a very recent paper of Fakcharoenphol and Talwar (see [4]), the KPR construction is slightly altered and, while properties 2 and 3 are preserved (up to constants), in property 1 the d -diameters become $O(r\Delta)$, improving the original $O(r^2\Delta)$. This has an immediate implication on our main result, yielding the following improvement:

$$\text{av}_{\text{line}}(d) \geq \Omega(r^{-2}) \cdot \text{av}(d).$$

Let us conclude this section with the following simple, yet interesting, observation:

Corollary 2.9 *Let (Y, Δ) be a special metric, e.g., a submetric of a shortest-path metric of a weighted planar graph. Then, for any metric d , $\text{av}_Y(d) \leq O(1) \cdot \text{av}_{\text{line}}$.*

Indeed, a good embedding of d into the line can be obtained by taking a good embedding of d into (Y, Δ) , and then embedding the resulting d' into the line as in the proof of Theorem 2.5, with only a constant loss in the average distance.

2.3 Constant Treewidth Metrics

Does the factor of $O(r^2)$ of (improved) Theorem 2.5 have the correct order of magnitude? In this section we address graphs of treewidth r , and show that for these graphs the true value of this factor is $O(\log r)$. In fact, the only property of such graphs which will be used is the existence of a balanced separator of size r . Instead of the KPR Lemma we use more ad-hoc methods, which nevertheless are of an independent interest.

We start with some preparatory results. Let d be the shortest-path metric of a (weighted) graph G on n vertices. Let $K \subset V$ be a subset of vertices whose removal results in connected components $\{S_i\}$ of sizes $\{n_i\}$, respectively. Let $\Delta = n^2 \text{av}(d)$ denote the sum of all the distances, and let Δ_X denote the sum of distances between the cross-pairs, i.e., all the pairs excluding those belonging to the same V_i . (A pair with $x \in K$ is never excluded!)

Lemma 2.10

$$\Delta \leq \max_i \frac{n + n_i}{n - n_i} \cdot \Delta_X.$$

Proof Let Δ_i be the sum of the distances between all the pairs of vertices in V_i . Arguing as in (2) of Sect. 1.1, we conclude that for every point p in G it holds

$$\Delta_i = \sum_{x,y \in S_i} d(x, y) \leq 2n_i \sum_{x \in S_i} d(p, x).$$

Thus,

$$\Delta_i \leq \frac{2n_i}{n - n_i} \sum_{p \notin S_i, x \in S_i} d(p, x).$$

Summing over all S_i 's, we get

$$\sum_i \Delta_i \leq \max_i \frac{2n_i}{n - n_i} \cdot \sum_i \sum_{p \notin S_i, x \in S_i} d(p, x) \leq \max_i \frac{2n_i}{n - n_i} \cdot \Delta_X.$$

Therefore,

$$\Delta - \Delta_X = \sum_i \Delta_i \leq \max_i \frac{2n_i}{n - n_i} \cdot \Delta_X,$$

and the lemma follows. □

Next, we need the following definition:

Definition 2.11 Let (S, d) be a finite metric space, let $K \subseteq S$ be a subset with the induced metric, and, finally, let $f : K \rightarrow \mathbb{R}$ be a nonexpanding embedding. Define the extensions f^+ and f^- of f to the entire S (to be called, respectively, the *rightmost* and the *leftmost* extension) by:

$$f^+(x) = \min_{y \in K} f(y) + d(y, x);$$

$$f^-(x) = \max_{y \in K} f(y) - d(y, x).$$

The fact that f is nonexpanding implies that both f^- and f^+ agree with f on K . To see that the extensions are nonexpanding, consider e.g., f^+ , and some two points $x_1, x_2 \in S$. Let $y_1, y_2 \in K$ be the two points in K that defined the values of $f^+(x_1)$ and $f^+(x_2)$, respectively. By definition of f^+ ,

$$f^+(x_1) - f^+(x_2) = [f(y_1) + d(y_1, x_1)] - [f(y_2) + d(y_2, x_2)]$$

$$\leq [f(y_2) + d(y_2, x_1)] - [f(y_2) + d(y_2, x_2)].$$

The last term is equal to $d(y_2, x_1) - d(y_2, x_2)$, and cannot exceed $d(x_1, x_2)$.

Observe that, as one would expect from the names, the rightmost extension f^+ maps every point $x \in S$ to its rightmost possible value among all nonexpanding extensions of $f|_K$, while f^- maps x to its leftmost possible value.

Definition 2.11 and Lemma 2.10 suggest the following plan for embedding into \mathbb{R} the shortest-path metric d of a graph G with a constant-size balanced separator K .

First, produce a suitable nonexpanding embedding $f : K \rightarrow \mathbb{R}$, or, rather, a suitable distribution over such embeddings. Then, for every connected component S_i of $G - K$, extend $f|_K$ randomly and independently from the others either as f^+ or as f^- . Since K is a separator, this results in a nonexpanding embedding of the entire $[n]$. Finally, trust your luck and try to verify that for any x, y not in the same S_i , the expected distance $E[|f^*(x) - f^*(y)|]$ is not far from the actual distance $d(x, y)$. The plan indeed works!

Theorem 2.12 *Let d be the shortest path metric of a (weighted) graph G of tree-width r . Then,*

$$\text{av}_{\text{line}}(d) \geq \Omega(1/\log r) \cdot \text{av}(d).$$

Proof Let $K \subseteq V[G] = [n]$ be a separator of G such that $|K| \leq r$, and all the n_i 's are at most $n/2$. We view K as a metric space equipped with the induced metric.

Recall that the standard construction of Bourgain [2] yields a distribution over nonexpanding embeddings $f : K \rightarrow \mathbb{R}$, such that for any two vertices $a, b \in K$, $E[d'(a, b)] \geq c_r d(a, b)$, where $c_r = \Omega(1/\log |K|) = \Omega(1/\log r)$.

The suitable distribution over the nonexpanding extensions of $f|_K$ will be defined as follows. With probability 0.5 take (randomly) one of the Bourgain's embeddings, and with probability 0.5 map the whole K to a single point. Extending every such $f|_K$ to V in a random manner according to the above plan, we obtain a distribution over embeddings $f : V \rightarrow \mathbb{R}$. We claim that for any x, y not in the same connected component of $G - K$, it holds that

$$E[|f(x) - f(y)|] \geq c_r/20 \cdot d(x, y) = \Omega(1/\log r) \cdot d(x, y).$$

In view of Lemma 2.10, this would imply the theorem.

Call x (respectively, y) *remote*, if $d(x, K) \geq c_r/5 \cdot d(x, y)$. Set $\delta = d(x, y)$.

Let us first consider the case when neither x nor y are remote. Let $p_x \in K$ be a vertex such that $d(x, p_x) < (c_r/5) \cdot \delta$; the vertex p_y is defined similarly. Observe that the distance $d(p_x, p_y)$ is at least $(1 - 2c_r/5) \cdot \delta$, and therefore the expected value of $|f(p_x) - f(p_y)|$ under Bourgain's embeddings is at least $(c_r - 2c_r^2/5) \cdot \delta$. Consequently, since with probability 0.5 one proceeds to choose a random Bourgain's embedding,

$$\begin{aligned} E[|f(x) - f(y)|] &\geq E[|f(p_x) - f(p_y)|] - E[|f(p_x) - f(x)|] - E[|f(p_y) - f(y)|] \\ &\geq 0.5 \cdot ((c_r - 2c_r^2/5) - c_r/5 - c_r/5) \cdot \delta \geq c_r/10 \cdot \delta, \end{aligned}$$

as required.

Consider now the second case, when w.l.o.g., x is remote. When K is mapped to a single point p , it holds

$$|f(x) - p| = d(x, K) \geq c_r/5 \cdot \delta,$$

and, independently, with probability 0.5 the images of x and y lie on the different sides of p . Since K is mapped to a single point with probability 0.5, we get

$$E[|f(x) - f(y)|] \geq (0.5 \cdot 0.5 \cdot c_r/5) \cdot \delta = c_r/20 \cdot \delta. \quad \square$$

Acknowledgements I am deeply grateful to Jirka Matoušek, Nati Linial and Yuval Rabani for enlightening discussions and helpful suggestions. Thanks also to Avner Magen.

References

1. Arora, S., Rao, S., Vazirani, V.: Expander flows, geometric embeddings, and graph partitioning. In: Proceedings of STOC 2004, pp. 222–231 (2004)
2. Bourgain, J.: On Lipschitz embeddings of finite metric spaces in Hilbert space. *Israel J. Math.* **52**(1–2), 46–52 (1985)
3. Chekuri, C., Gupta, A., Newman, I., Rabinovich, Y., Sinclair, A.: Embedding k -outerplanar graphs into ℓ_1 . In: Proceedings of the 14th Annual ACM Symposium on Discrete Algorithms, pp. 527–536 (2003)
4. Fakcharoenphol, J., Talwar, K.: An improved decomposition theorem for graphs excluding a fixed minor. In: Proceedings of RANDOM-APPROX 2003, pp. 36–46 (2003)
5. Fakcharoenphol, J., Rao, S., Talwar, K.: A tight bound on approximation arbitrary matrix by tree matrix. *J. Comput. Syst. Sci.* **69**(3), 485–497 (2004)
6. Gromov, M.: *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhäuser, Boston (1999)
7. Gupta, A., Newman, I., Rabinovich, Y., Sinclair, A.: Cuts, trees and ℓ_1 -embeddings of graphs. In: Proceedings of the 40th Annual IEEE Symposium on Foundations of CS, pp. 399–408 (1999)
8. Gusfield, D.: *Algorithms on Strings, Trees, and Sequences*. Cambridge University Press, Cambridge (1997)
9. Indyk, P.: Algorithmic aspects of geometric embeddings. In: Proceedings of the 42th Annual IEEE Symposium on Foundations of Computer Science (2001)
10. Khot, S., Vishnoi, N.K.: The unique games conjecture, Integrality gap for cut problems and embeddability of negative type metrics into L_1 . In: Proc. FOCS 2005, pp. 53–62 (2005)
11. Klein, P., Plotkin, S.A., Rao, S.B.: Excluded minors, network decomposition, and multicommodity flow. In: Proceedings of the 25th Annual ACM Symposium on Theory of Computing, pp. 682–690 (1993)
12. Lee, J.R., Naor, A.: Embedding the diamond graph in L_p , and dimension reduction in L_1 . Manuscript, June 2003. Submitted to GAFA
13. Linial, N., London, E., Rabinovich, Y.: The geometry of graphs and some of its algorithmic applications. *Combinatorica* **15**(2), 215–245 (1995)
14. Matoušek, J.: *Lectures on Discrete Geometry*, vol. 212. Springer, Berlin (2002)
15. Matoušek, J., Rabinovich, Y.: On dominated l_1 metrics. *Israel J. Math.* **123**, 285–301 (2001)
16. Milman, V., Schechtman, G.: *Asymptotic Theory of Finite Dimensional Spaces*. Lecture Notes in Mathematics, vol. 1200. Springer, Berlin (1986)
17. Newman, I., Rabinovich, Y.: A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete Comput. Geom.* **29**(1), 77–81 (1998)
18. Rao, S.: Small distortion and volume preserving of planar and Euclidean metrics. Preliminary version in: 15th Annual ACM Symposium on Computational Geometry, pp. 300–306 (1999)